Həyat elmləri və biotibb jurnalı, cild 74, №2, səh. 5-12 (2019)

Atom yerdəyişmə parametrlərindən istifadə etməklə Zülalların Verilənlər Bazasında olan quruluşların yoxlanılması

R.Ç. Məsməliyeva

AMEA Molekulyar Biologiya və Biotexnologiyalar İnstitutu, İzzət Nəbiyev, 11, Bakı AZ1073, Azərbaycan;
E-mail: r.masmaliyeva@imbb.science.az
Çapa qəbul edilmişdir: 28.12.2019

Hal-hazırda Zülalların Verilənlər Bazasında (ZVB) 156 mindən çox makromolekulyar quruluş vardır. Əgər ZVB yaradılandan sonra ilk dövrlərdə mümkün qədər çox quruluşların alınması prioritet idisə, indi əsas məqsəd müxtəlif bioloji prosesləri başa düşmək üçün məqsədə uyğun olan molekulların quruluşlarının öyrənilməsidir. Bu isə molekulların üçölçülü quruluşlarının yoxlanılması və yaxşılaşdırılması probleminin əhəmiyyətini daha da artırır. Yeni yoxlama metodlarının yaradılması və inkişaf etdirilməsi vacibdir. Klassik yoxlama metodları zülal molekullarının fiziki və kimyəvi xüsusiyyətlərini, amin turşuların həndəsəsini nəzərə alır. Tədqiqatlarımızda göstəririk ki, atom yerdəyişmə parametrləri (AYP) də bu məqsədlə istifadə oluna bilər. AYP atomun öz pozisiyasından yerini dəyişməsini ifadə edir. Onların statistik modelinin inkişaf etdirilməsi makromolekulyar quruluşların yoxlanılması və yaxşılaşdırılması üçün əhəmiyyətli vasitə ola bilər. Bu məqalə temperatur faktorunun statistik analizinin bütün ZVB-yə tətbiqinə və sözügedən yoxlama metodunun məlum quruluşların yaxşılaşdırılması üçün mümkün istifdəsinə həsr olunmuşdur.

Açar sözlər: *Atom yerdəyişmə parametrləri, makromolekulyar saflaşdırma, validasiya*

GİRİŞ

Hal-hazırda Zülalların Verilənlər Bazasında 156 mindən çox quruluş vardır və hər il bu bazaya təxminən 10 min yeni quruluş əlavə olunur. (Berman et al., 2002). Əgər ZVB yaradılandan sonra ilk dövrlərdə prioritet mümkün qədər çox quruluşların alınması idisə, indi, əsas məqsəd konkret quruluşların öyrənilməsidir. Bu isə molekulların üçölçülü quruluşlarının yoxlanılması və yaxşılaşdırılması probleminin əhəmiyyətini daha da artırır. Əvvəlki məqaləmiz (Masmaliyeva and Murshudov, 2019) atom yerdəyişmə parametrlərindən (AYP) (Trueblood et al., 1996) biri olan izotropik temperatur faktorunun (və ya B faktor) statistik analizinin makromolekulyar quruluşlarının yoxlanılması üçün istifadəsində əhəmiyyətinə həsr edilmişdir. Zülal quruluşlarının yoxlanılmasına həsr olunan məqalələr çox olsa da, AYP-nin analizi cəmi bir neçə tədqiqat işində əks olunmuşdur (Dauter et al., 2006; Merritt, 2011, Read, 2011; Schneider, 2014, Carugo, 2018; Grosse-Kunstleve, 2002). Adətən molekulyar quruluşların yoxlanılması üçün atomun nisbi pozisiyasından istifadə olunur. AYP-i atomların pozisiyalarının dəqiqliyini ifadə etdiyindən, onlardan da quruluş modellərinin yoxlnılmasında istifadə etmək məqsədyönlüdür. AYP-i quruluş modellərinin vacib hissəsi olub, atomların öz pozisiyasından yayınmasını ifadə edir. B faktorun saflaşdırılması eksperimental verilənlər və model arasındakı uyğunluğu yüksəldir. Temperatur faktorunun mütləq qiyməti ilə problem ondadır ki, onun bütün qiymətlərinə eyni bir ədədi əlavə etsək/çıxsaq, orta qiymət uyğun olaraq dəyişəcək, lakin varians dəyişmir. Elektron paylanmalarının Fürye analizindən məlumdur ki, makromolekulyar quruluşların təsvirlərinin həddən artıq kəskinkəşdirilməsi və ya yayılması mənasızdır: Fürye sıralarının kəsilməsi effekti həddən artıq çox olduqda küyün qiyməti güclənərək siqnalın görünməsinə imkan verməyəcək. Ona görə də bu hallar arzuedilməz hesab olunur. Bu da saflaşdırma alqoritmlərində B qiymətlərinə əsaslanan məhdudlaşdırmaların tətbiqində əlverişsizdir. Bu səbəbdən də B faktorun paylanmasından istifadə etmək daha məqsədəuyğundur. İzotropik temperatur faktorunun paylanması Sürüşən tərs qamma paylanmasına uyğun gəlir (Masmaliyeva and

Murshudov, 2019). Bu paylanmanın parametrlərinin qiymətləndirilməsi və qrafik təsviri zülal molekullarının üçölçülü modellərindəki səhvlərin aşkarlanmasına imkan verir. Bu məqalə temperatur faktorunun statistik analizinin bütün ZVB-yə tətbiqinə və hazırlanmış yoxlama metodunun məlum quruluşların yoxlanılması üçün mümkün istifədəsinə həsr olunub.

MATERİAL VƏ METODLAR

ZVB-dakı quruluşlar yüklənmiş və Refmac (Murshudov et al., 2011) proqramı ilə yenidən saflaşdırılma aparılmışdır. R faktorun (Morris et al., 1992) qiyməti 0.30 və yuxarı olan quruluşlar analizlərdən kənarlaşdırılmışdır. Belə ki, bu quruluşlarda model və eksperimental verilənlərin uyğunluğu zəifdir. Yenidən saflaşdırma verilənlərin obyektivliyi baxımından vacibdir. Çünkü, baxmayaraq ki ZVB-dakı quruluşların bir hissəsi Refmac proqramı ilə saflaşdırılsa da, digər proqramlarla saflaşdırılan quruluşlar da vardır. Təcrübə göstərir ki, müasir üsullardan istifadə edərək yenidən saflaşdırma verilənlərin yoxlanılmasından düzgün nəticə almaq üçün zəruri prosedurdur.

Sürüşən tərs qamma paylanmasının parametrləri bunlardır: forma parametri (α), miqyas parametri (β) və sürüşmə addımı parametri (B_0):

$$\boldsymbol{P(B; B_0; \alpha; \beta) = \frac{\beta^\alpha}{\Gamma(\alpha)}(B - B_0)^{-\alpha-1} exp\left(\frac{-\beta}{B-B_0}\right)} \quad \textbf{(1)}$$

İlk iki momentdən istifadə etməklə (momentlər üsulu) parametrlərdən ikisini hesablamaq olar (Stuart et al., 1999):

$$\boldsymbol{\langle B - B_0 \rangle = \frac{\beta}{\alpha - 1}}$$

$$\boldsymbol{var(B - B_0) = var(B) = \frac{\beta^2}{(\alpha-1)^2(\alpha-2)}} \quad \textbf{(2)}$$

Üçüncü parametri isə ya üçüncü momentdən və ya minimum B faktordan istifadə edərək tapmaq olar. Alınan parameterlər sonra Maksimal mümkünlük metodu ilə daha da yaxşılaşdırılır.

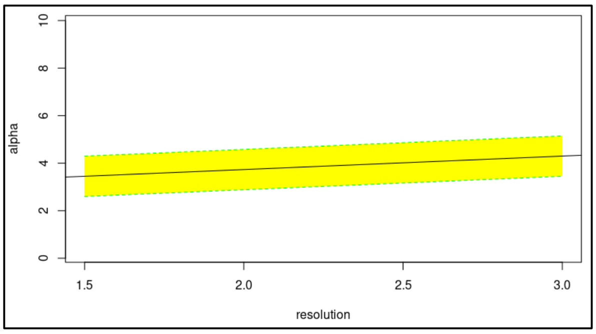
Atom yerdəyişmə parametrlərinin paylanmasını Sürüşən tərs qamma paylanması ilə modelləşdirdikdən və bütün ZVB üçün parametrləmə apardıqdan sonra *R* statistika paketinin (R core team, 2014) köməyi ilə Proqnozlaşdırma intervalı hesablanmışdır. Bu intervaldan kənar qiymətlər göstərən quruluşlar daha dərindən nəzərdən keçirilməlidir. Proqnozlaşdırma intervalı reqressiya analizində proqnozlaşdırma üçün istifadə olunan Etibarlılıq intervalı növüdür (Stuart et al., 1999). Bu elə qiymətlər parçasıdır ki, bu, verilən modelə uyğun yeni müşahidələrin proqnozlanmasına imkan verir. Proqnozlaşdırma intervalının Etibarlılıq intervalından əsas fərqi odur ki, Etibalılıq intervalı çoxluğun öz parametrləri ilə əlaqədar olduğu halda, Proqnozlaşdırma intervalı gələcək müşahidələrin qiymətlərinin düşəcəyi parçanı ifadə edir.

Bu parametrlərin hesablanması və B faktorun paylanmasının qrafiki təsvirinin qurulması üçün ToBvalid.py proqram təminatı inkişaf etdirilmiş və istifadə edilmişdir (Masmaliyeva and Murshudov, 2019). Eləcə də, Sürüşən tərs qamma paylanmasının ümumi statistikası aparılmışdır: minimal qiymət, maksimal qiymət, orta qiymət, median, varians, meyillilik, kurtozis, 1-ci kvartil, 3-cü kvartil. Kənaraçıxmaların aradan qaldırılması üçün İnterkvartil məsafələr və ya Tukey metodundan (Tukey et al., 1977; McGill, 1978, Masmaliyeva and Murshudov, 2018) və Tərs qamma üçün Kvantaillər funksiyasından (Jones et al., 2019) istifadə olunmuşdur. Tukey metodunda əmsal 3 istifadə edilmişdir. Bu əmsal temperatur faktorunun paylanmasında kənaraçıxmaların aradan qaldırılması üçün daha əlverişlidir (Masmaliyeva and Murshudov, 2018). Kvantaillar funksiyasında persentil nöqtələri 0.05 və 0.95 seçilmişdir. Qeyd etmək lazımdır ki, Tərs qamma paylanması sağ meyillilikli asimmetrik paylanma olduğuna görə Tukey metodu əksər hallarda ancaq sağ kənaraçıxmaları müəyyən edir. Lakin aşırı yayılmanı ifadə edən böyük qiymətlərlə yanaşı, B faktorun öz mühitindən kəskin fərqlənən çox kiçik qiyməylərini müəyyən etmək yanlış yerləşdirilmiş və ya ağır atomların müəyyənləşdirilməsində əhəmiyyət kəsb edir. Python *SciPy.stats* (Jones et al., 2019) kitabxamasının *invgamma* funksiyasının kvantaillar metodu bu kənaraçıxmaları da müəyyən etəməyə imkan verir və Tərs qamma paylanmasının xüsusiyyətlərini nəzərə alır.

NƏTİCƏLƏR VƏ ONLARIN MÜZAKİRƏSİ

α forma parametrinin proqnozlaşdırma intervalı. Atom yerdəyişmə parametrlərinin Sürüşən tərs qamma paylanması modelini bütün ZVB-yə

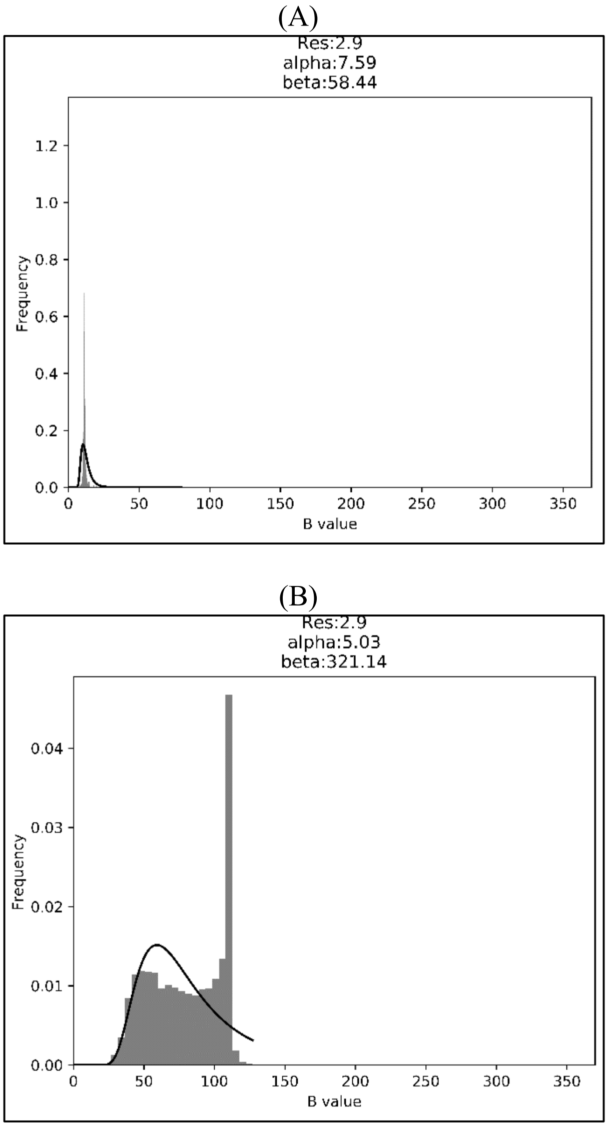
tətbiq edəndən sonra ZVB-dəki quruluşların əksəriyyətinin bu modelə uyğun gəldiyi məlum olmuşdur. ZVB-dəki bütün quruluşların α parametrlərindən istifadə etməklə Proqnozlaşdırma intervalı hesablanmışdır (Şəkil 1). Tənlik 2-dən görünür ki, yalnız α parametrinin qiyməti 1-dən böyük olduğu halda 1-ci momentin və 2-dən böyük olduğu halda 2-ci momentin mənası vardır. Proqnozlaşdırma intervalının qrafik təsvirindən görünür ki, bu, zülal quruluşları üçün doğrudur.



Səkil 1. α parametrinin ayırdetməyə görə Proqnozlaşdırma intervalı.

Saflaşdırma metodunun seçilməsi. B faktorun paylanmasının Sürüşən tərs qamma paylanmasından istifadə etməklə modelləşdirilməsinin saflaşdırma metodunun seçilməsində də əhəmiyyətli rolu vardır. α və β parametrlərinin çox böyük qiymətləri, xüsusilə də α parametrinin Proqnozlaşdırma intervalı xaricindəki qiymətləri tətbiq edilən saflaşdırma prosesinin qeyri-optimal olmasının nəticəsində meydana çıxa bilər. Bu halda yenidən saflaşdırmanın tətbiq edilməsi zəruridir. Buna misal ola bilən quruluşlardan biri ZVB-də 1OB1 identifikatorlu quruluş olan Fab kompleksin kristal strukturudur (Pizzaro et al., 2003). Qeyd etmək lazımdır ki, bu kompleks *Plasmodium falciparum*-dan alınmış və malyariya xəstəliyi vaksininin hazırlanmasında istifadə edilmək üçün namizəddir. Şəkil 2-də 1OB1 quruluşunun yenidən saflaşdırmadan əvvəl (A) və sükuta görə yenidən saflaşdırmadan sonra (B) temperatur faktorunun paylanmasının histoqramı və Sürüşən tərs qamma paylanması təsvir edilmişdir.

Şəkil 2-dən göründüyü kimi, bu paylanmalar Tərs qamma paylanmasına uyğun gəlmir, və α parametri Proqnozlaşdırma intervalına uyğun deyildir. Yenidən saflaşdırmanın dövrlərinin sayını 20 (sükuta görə Refmac proqramında 5 dövr təyin edlib) təyin etdikdən və yayılma (25 Å) tətbiq edəndən sonra, B faktorun paylanması Tərs qammaya uyğun gəlir, α parametrinin qiyməti Proqnozlaşdırma intervalına daxil olur və həmçinin R faktorun qiyməti aşağı düşür (Şəkil 3, Cədvəl 1).

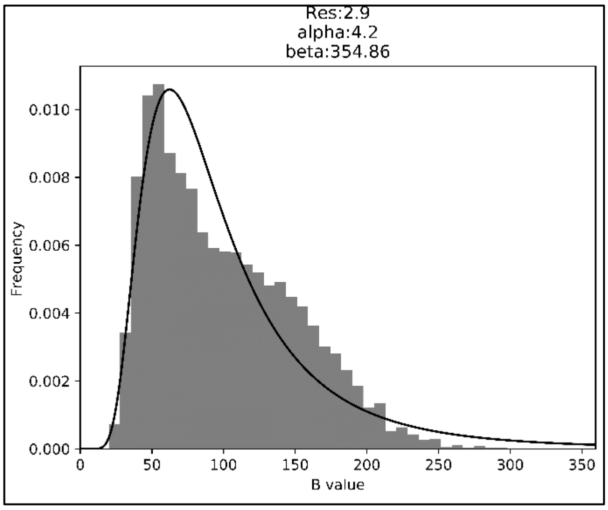


Şəkil 2. 1OB1 quruluşunun yenidən saflaşdırmadan əvvəl (A) və sükuta görə yenidən saflaşdırmadan sonra (B) temperatur faktorunun paylanmasının histoqramı və Sürüşən tərs qamma paylanması.

Cədvəl 1. 1OB1 quruluşunun B faktor paylanmasının tərs qamma parametrlərinə saflaşdırma üsulunun təsiri.

	α	**β**	**R faktor**
ZVB-dən yüklənmiş quruluş	7.59	58.44	0.258[1]
Standart yenidən saflaşdırma	5.03	321.14	0.28041
Sonrakı yenidən saflaşdırma	4.2	354.86	0.22507

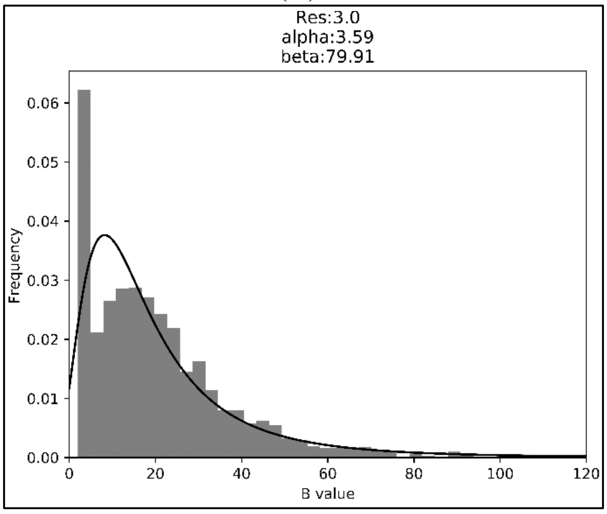
[1]R faktorun ZVB-də verilmiş qiyməti



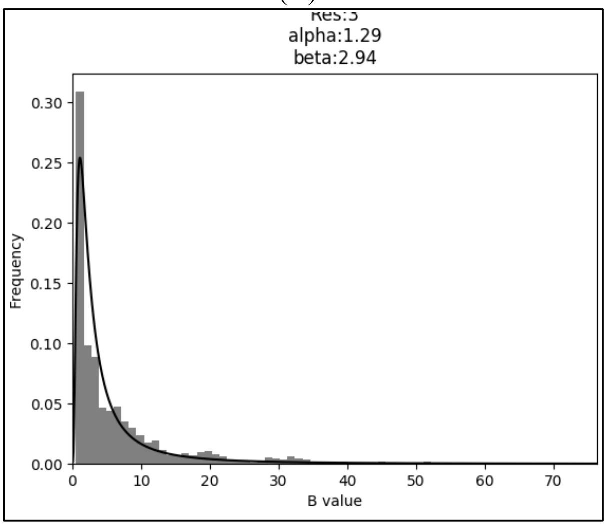
Şəkil 3. 1OB1 quruluşunun istifadəçi tərəfindən müəyyən edilən yenidən saflaşdırmadan (20 dövr yenidən saflaşdırma və 25 Å yayılma) sonra temperatur faktorunun paylanmasının histoqramı və Sürüşən tərs qamma paylanması.

Aşırı kəskinləşdirilmiş hal. Verilmiş paylanmada B faktorun minimal qiymətlərinin tezliyi yuxarı olduqda paylanma sola sürüşmüş olarsa, demək bu quruluş saflaşdırmadan qabaq aşırı kəskinləşdirilmişdir. Belə halda Sürüşən tərs qamma paylanmasının parametrləri qeyri-tipik qiymətlər alır: B_0 sürüşmə addımı parametri sıfra yaxınlaşır və hətta mənfi qiymətlər də ala bilir, α parametri proqnozlaşdırma intervalının aşağı sərhəddindən kiçik olur. Bütün ZVB-dəki quruluşlardan 3252 ədəd $B_0 \leq 0$ halı vardır. Ümumilikdə ZVB-də 5868 aşırı kəskinləşdirilmiş quruluş müəyyən edilmişdir. Bu quruluşlara aid nümunələrdən biri ZVB-də 1AOW identifikatorlu kalsium/fosfolipid birləşdirici zülal olan Anneksin IV-ün quruluşudur. Burada B faktorun minimal qiyməti və ona yaxın qiymətlərin yüksək tezliyini və paylanmanın sola sürüşmüş histoqramını müşahidə etmək olar (Şəkil 4). Yenidən saflaşdırma zamanı yayılma (25 Å) tətbiq edəndən sonra B faktorun paylanması Sürüşən tərs qamma paylanmasına uyğun gəlir (Şəkil 5, Cədvəl 2).

(A)



(B)

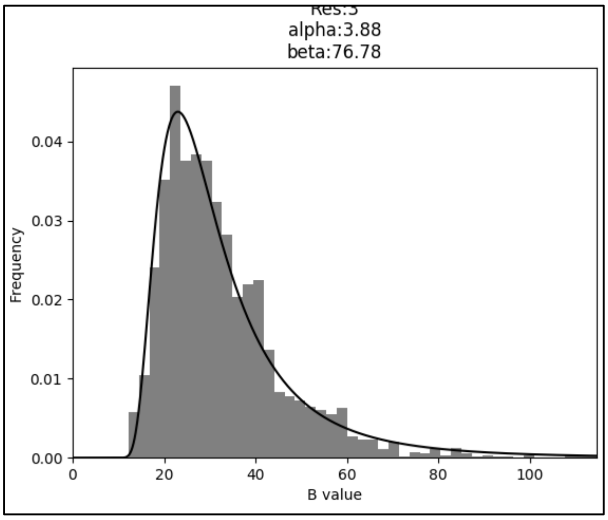


Şəkil 4. 1AOW quruluşunun yenidən saflaşdırmadan əvvəl (A) və sükuta görə yenidən saflaşdırmadan sonra (B) temperatur faktorunun paylanmasının histoqramı və Sürüşən tərs qamma paylanması.

Cədvəl 2. 1AOW quruluşunun B faktor paylanmasının tərs qamma parametrlərinə saflaşdırma üsulunun təsiri

	α	β	B_0	R faktor
ZVB-dən yüklənmiş quruluş	3.59	79.91	-9.21	0.19[2]
Standart yenidən saflaşdırma	1.29	2.91	-0.19	0.24
Sonrakı yenidən saflaşdırma	3.73	59.85	8.81	0.23

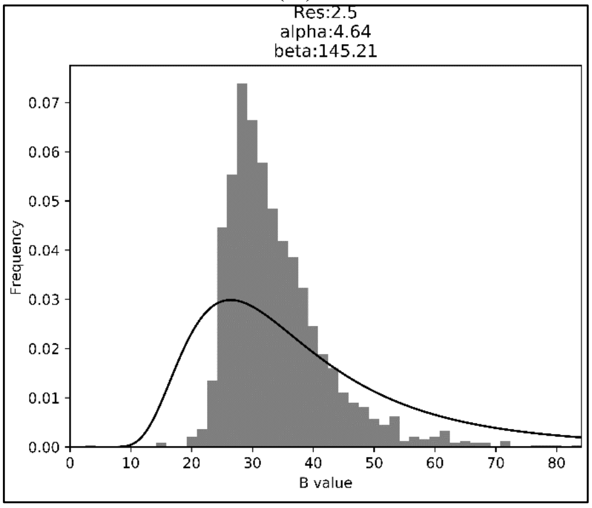
[2]R faktorun ZVB-də verilmiş qiyməti



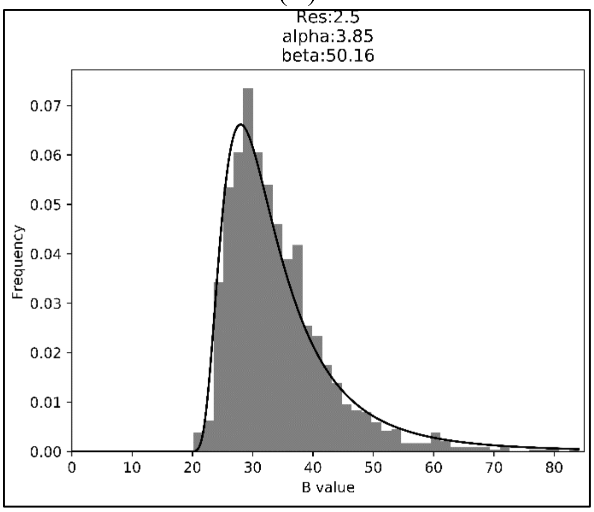
Şəkil 5. 1AOW quruluşunun istifadəçi tərəfindən müəyyən edilən (25 Å yayılma) yenidən saflaşdırmadan sonra temperatur faktorunun paylanmasının histoqramı və Sürüşən tərs qamma paylanması.

Sürüşən tərs qamma paylanmasında sol quyruq müşahidə edilən hal. Tərs qamma paylanması müsbət asimmetriyalı, sağ quyruqlu müsbət ehtimal paylanmasıdır. Temperatur faktorunun Sürüşən tərs qamma paylanması kimi modelləşdirilməsi alqoritmini ZVB-dəki bütün zülal quruluşlara tətbiq etdikdən sonra müəyyən edilmişdir ki, 1469 quruluşda B faktorun paylanmasında sol quyruq müşahidə olunur. Ümumiyyətlə bu hal onu göstərir ki, quruluş modelində yanlış modelləşdirilmiş atomlar və ya ağır atomlar ola bilər. Bu hala aid bir nümunəni ZVB-də 1ZNZ identifikatorlu QTF ilə kompleksdə Quanilat kinazanın (Hible, 2005) üçölçülü quruluş modelinin misalında nəzərdən keçirək. Sol quyruqluluq halının meydana çıxması ona gətirib çıxarır ki, histoqram və tərs qamma paylanması bir-birinə uyğun gəlmir və bu kənaraçıxmaların ehtimal paylanmasına təsiri nəticəsində hesablanmış parametrlər B faktorun paylanmasına tam uyğun gəlməyə bilər (Şəkil 6 A). Tərs qamma üçün kvantaillar metodundan istifadə etməklə sol quyruqluluq əmələ gətirən kənaraçıxmaları müəyyən etmək olar. Bu kənaraçıxmalar aradan qaldırıldıqdan sonra B faktorun histoqramı və hesablanan parametrlər uyğun tərs qamma paylanması üst-üstə düşür (Şəkil 6 B).

(A)



(B)



Şəkil 6. Bir ZNZ identifikatorlu zülalın quruluş modelində B faktorun paylanmasından sol quyruqluluq əmələ gətirən kənaraçıxmaları aradan qaldırmazdan əvvəl (A) və aradan qaldırdıqdan sonra (B) Sürüşən tərs qamma paylanmasının necə dəyişdiyinin təsviri.

Cədvəl 3. Bir ZNZ quruluşunun B faktor paylanmasının parametrlərinə kənaraçıxmaların aradan qaldırılmasının təsiri.

	α	β	B_0
Kənaraçıxmalar ilə birlikdə	4.64	145.21	0.68
Kənaraçıxmalar aradan qaldırıldıqdan sonra	3.85	50.16	17.64

Parametrlərin necə dəyişməsi Cədvəl 3-də verilmişdir.

Beləliklə, atom yerdəyişmə parametrlərinin statistik modeli zülal quruluşlarının yoxlanılmasında əhəmiyyətli vasitədir. Aşırı kəskinləşdirmə yaxud da yayılma hallarının mümkünlüyü Temperatur faktorunun mütləq qiymətini mənasız edir. B faktorun mütləq qiyməti əvəzinə, onun paylanmasının istifadə edilməsi daha məqsədəuyğundur. Nadir hallarda, bəzi quruluşlar üçün temperatur faktorunun statistik paylanması Sürüşən tərs qamma paylanmasına uyğun gəlməyə bilər. Bu onu göstərir ki, belə quruluş modeli daha dərindən analiz edilməlidir. B faktorun paylanmasının Tərs qammaya uyğun gəlməməsinin bir neçə səbəbi ola bilər. Bunlardan, çox yüksək aşırı-kəskinləşdirmə və ya bulanıqlılıq, bəzi quruluşlar üçün isə tvinninq halları, həmçinin, saflaşdırma zamanı qeyri-kristalloqrafik simmetriya hissəsində məhdudlaşdırmaların (sərt, qlobal, lokal) düzgün seçilməməsi ola bilər. Ümumilikdə isə B faktorun belə modelləşdirilməsi ZVB-dəki quruluşlar üçün özünü doğrultdur. Temperatur faktorunun statistik paylanmasının parametrlənməsi saflaşdırmadan əvvəl aşırı kəskinləşdirmə yaxud da yayılmanın olmasını müəyyən etməyə imkan verir. Gələcəkdə Temperatur faktorunun paylanmasının aşağı ayırdetməli quruluşların saflaşdırılmasına tətbiq edilməsi mümkündür. Bu metod, saflaşdırmadan öncə kəskinləşdirmə dərəcəsinin düzgün təyin edilməsi üçün xüsusilə əhəmiyyətli ola bilər. Temperatur faktorunun ayırdetmə ilə əlaqəli nəzərdən keçirilməsi həmişə yaxsı ideyadır. Gələcək tədqiqatlar həmçinin B faktorun qiyməti və ayırdetmədən asılı olaraq, nöqtəvi atomun ətrafında elektron buludunun hündürlüyündən (Masmaliyeva and Murshudov, 2019) istifadə etməklə, B faktorunun qiyməti öz yaxın ətrafından kəskin fərqlənən yanlış modelləşdirilmiş atomların müəyyənləşdirilməsinə istiqamətlənmiş alqoritmin inkişaf etdirilməsinə həsr ediləcəkdir.

MİNNƏTDARLIQ

Tədqiqat işi Azərbaycan Milli Elmlər Akademiyası Rəyasət Heyətinin 5/9 nömrəli qərarı əsasında maliyyələşən qrant layihəsi çərçivəsində həyata keçirilmişdir.

ƏDƏBİYYAT

Berman H.M., Battistuz T., Bhat T.N., Bluhm W.F., Bourne P.E., Burkhardt K., Feng Z., Gilliland G.L., Iype L., Jain S., Fagan P., Marvin J., Padilla D., Ravichandran V., Schneider B., Thanki N., Weissig H., Westbrook J.D., Zardecki C. (2002) The Protein Data Bank. *Acta Cryst.,* **D58:** 899-907.

Carugo O. (2018) How large B-factors can be in protein crystal structures. *BMC bioinformatics,* **19(1):** 61.

Dauter Z., Murshudov G.N., Wilson K.S. (2006) International Tables for Crystallography. **Vol. F:** 393-402. *Chester: International Union of Crystallography.*

Grosse-Kunstleve, R.W., Adams P.D. (2002) On the handling of atomic anisotropic displacement parameters. *J. Appl. Cryst.,* **35:** 477-480.

Hible G., Christova P., Renault L., Andrew E.S., Girard T.E., Munier-Lehmann H., Cherfils J. (2005) Unique GMP-binding site in *Mycobacterium tuberculosis* guanosine monophosphate kinase. *Proteins,* **62(2):** 489-500

Jones E., Oliphant E., Peterson P., et al. (2019) *SciPy: Open Source Scientific Tools for Python, 2001; http://www.scipy.org/*

Read R.J., Adams P.D., Arendall W.B. 3rd, Brunger A.T., Emsley P., Joosten R.P., Kleywegt G.J., Krissinel E.B., Lütteke T., Otwinowski Z., Perrakis A., Richardson J., Sheffler W., Smith J., Tickle I., Vriend G., Zwart P. (2011) A new generation of crystallographic validation tools for the protein data bank. *Structure,* **19:** 1395–1412.

Masmaliyeva R., Murshudov G.N. (2017) Refinement and validation of macromolecular structures. *Transactions of the Institute of Molecular Biology & Biotechnologies of ANAS,* **1:** 80-93.

Masmaliyeva R., Murshudov G.N. (2018) Outlier detection in atomic temperature factor - B value distribution. *Proceedings of ANAS (Biological and Medical Sciences),* **73(2):** 25-32.

Masmaliyeva R.C., Murshudov G.N. (2019) Analysis and validation of macromolecular B values. *Acta Cryst.,* **D75:** 505-518.

Merritt E.A. (2011) Some Beq are more equivalent than others. *Acta Cryst.,* **A67:** 512–516.

McGill R., Tukey John W., Larsen Wayne A. (1978) Variations of box plots. *The American Statistician.,* **32(1):** 12–16.

Morris A.L., MacArthur M.W., Hutchinson E.G., Thornton J.M. (1992) Stereochemical quality of protein structure coordinates. *Proteins,* **12 (4):** 345–64.

Murshudov G.N., Skubák P., Lebedev A.A., Pannu N.S., Steiner R.A., Nicholls R.A., Winn M.D., Long F., Vagin A.A. (2011) REFMAC5 for the refinement of macromolecular crystal structures. *Acta Cryst.,* **D67:** 355–367.

Pizarro J.C., Chitarra V., Verger D., Holm I., Petres S., Dartville S., Nato F., Longacre S., Bentley G.A. (2003) Crystal structure of a Fab complex whith Plasmodium falciparum MSP1-19.J. Mol.Biol., **328:** 1091.

R Core Team (2014) R: A Language and Environment for Statistical Computing. Vienna: R Foundation for Statistical Computing. *http://www.R-project.org.*

Read R.J., Adams P.D., Arendall W.B., Brunger A.T., Emsley P., Joosten R.P., Kleywegt G.J., Krissinel E.B., Lütteke T., Otwinowski Z., Perrakis A., Richardson J.S., Sheffler W.H., Smith J.L., Tickle I.J., Vriend G., Zwart P.H. (2011) A new generation of crystallographic validation tools for the protein data bank. *Structure,* **12; 19(10):**1395-412.

Schneider B., Gelly J.C., de Brevern A.G., Černý J. (2014) Local dynamics of proteins and DNA evaluated from crystallographic B factors. *Acta Cryst., Section D, Biological crystallography,* **70(Pt 9):** 2413–2419.

Stuart A., Ord K., Arnold S. (1999) Kendall's Advanced Theory of Statistics, ***2A.*** *London: Edward Arnold.*

Trueblood K.N., Burgi H.-B., Burzlaff H., Dunitz J.D., Gramaccioli C.M., Schulz H.H., Shmueli U., Abrahams S.C. (1996) Atomic Dispacement Parameter Nomenclature. Report of a Subcommittee on Atomic Displacement Parameter Nomenclature. *Acta Cryst.,* **A52: 770-781.**

Tukey J.W. (1977). Exploratory data analysis. *Addison-Wesley Series in Behavioral Science - Quantitative Methods,* 711.

Анализ структур Базы Данных Белков с использованием параметров смещения атомов

Р.Ч. Масмалиева

Институт молекулярной биологии и биотехнологий НАН Азербайджана, Баку, Азербайджан

В настоящее время в Базе Данных Белков (БДБ) имеется более 156 000 макромолекулярных структур. Если на первых этапах развития структурной биологии приоритетной задачей являлось получение как можно большего числа структур, то в настоящее время основная задача заключается, в разъяснении сути различных биологических процессов путем целенаправленного изучения строения молекул. В связи с этим проверка и уточнение трехмерных макромолекулярных структур приобретает все более важное значение, т.е. необходима разработка новых программ проверки. Классические методы валидации основаны на физических и химических особенностях макромолекул, а также на геометрии аминокислот. В наших исследованиях мы описываем метод проверки, основанные на использования параметров смещения атома (ПСА). ПСА иллюстрирует отклонение атома от

его основного положения. Разработанный метод статистического анализа распределения ПСА может быть использован в качестве эффективного инструмента валидации. Данная статья посвящена применению статистического анализа ПСА к ББД и возможному использованию этих методов для макромолекулярной валидации.

Ключевые слова: *Параметры смещения атома, макромолекулярное уточнение, валидация*

Analysis of structures from Protein Data Bank using atom displacement parameters

R.C. Masmaliyeva

Institute of Molecular Biology and Biotechnologies, Azerbaijan National Academy of Sciences, Baku, Azerbaijan

Currently there are more than 156,000 macromolecular structures in Protein Data Bank (PDB). If at first times of development of Structural biology the priority purpose was solving as many structures as possible, currently the main purpose is solving structures, which allow understanding different biological processes. This is why validation and refinement of macromolecular structures become more and more important. New validation tools need to be developed. Classic methods of validation are based on physical and chemical features of macromolecules, as well as the geometry of amino acids. In our researches, we describe a method of validation by using Atom Displacement Parameters (ADP). ADP illustrates the variation of atom from its main position. Development of the method of statistical analysis of ADP distribution may be used as an effective validation tool. This contribution is devoted to application of statistical analysis of ADPs to PDB and possible use of this method for macromolecular validation.

Keywords: *Atom displacement parameters, macromolecular refinement, validation*